

ABSTRACT

A data format for transmitting replica data between a source system and a destination system provides various standalone headers and data following headers that identify aspects of the data stream in a generalized manner. More particularly, the header layout includes data following type headers structuring file names so they interoperate with systems of different language settings, character representations, or alternate means of naming data. The header layout also uses an extended attribute field in which a number of different platform/operating system-specific attributes (e.g., ACLs, streams, etc.) are identified and associated with particular named files in the follow-on data stream. An additional space in the header is set aside for carrying the names in association with the extended attributes. On the destination storage system, a set of hidden metadata directories ("metadirs") are provided for storing extended attributes in association with any and all files on the destination. The extended attributes are appended to each file, as applicable, via an NT stream. The metadir set contains (a) a permanent metadir which stores active copies of the files, available for restoration to the source; (b) a purgatory metadir, to which the files and streams of the permanent metadir are transferred during a update operation (e.g. receiving changed files from the source); and (c) a new metadir to which changed files with new extended attributes, if any, are written. The format also supports a standalone header that instructs the destination to undo a backed-up changed file in the event that the source determines that the file was modified in the midst on a backup procedure and transmission to the destination.

A system for transmitting replica data between a source system and a destination system provides various standalone headers and data following headers that identify aspects of the data stream in a generalized manner. On the destination storage system, a set of hidden metadata directories ("metadirs") are provided for storing extended attributes in association with files on the destination. The system also supports a standalone header that instructs the destination to undo a backed-up changed file in the event that the source

determines that the file was modified in the midst of a backup procedure and transmission to the destination.